

TD math-SES n° 5: à propos des « strobiloïdes »: des histogrammes aux courbes, du discret au continu ...

étude des inégalités de répartition des revenus des ménages à l'aide de graphiques. Les données, tableaux et graphiques proviennent de l'étude de l'INSEE : « les revenus et patrimoines des ménages, édition 2006 » d'après l'enquête 2003-2004.

I- Un graphique de base : le regroupement par classe des données, histogramme des effectifs et cumulés croissants.

Les données collectées lors d'une enquête sont regroupées en vue d'une étude statistique pour être analysées et commentées.

Les données de la série sont regroupées par classes dans un tableau des effectifs ou des fréquences puis représentées sur un graphique appelé histogramme en vue d'une interprétation.

Dans un histogramme, l'aire de chaque baton est proportionnelle à l'effectif de chaque classe.

Ainsi l'aire totale de tous les rectangles est égale à : _____

Quelle est la largeur de chacune des classes dans le graphique ci-contre ?

On l'appelle aussi le « pas » de l'histogramme.

Remarque :

Dans un histogramme ordinaire, le rectangle est construit sur toute la largeur de la classe mais l'Insee a ici voulu mettre en relation deux séries sur un même graphique ...

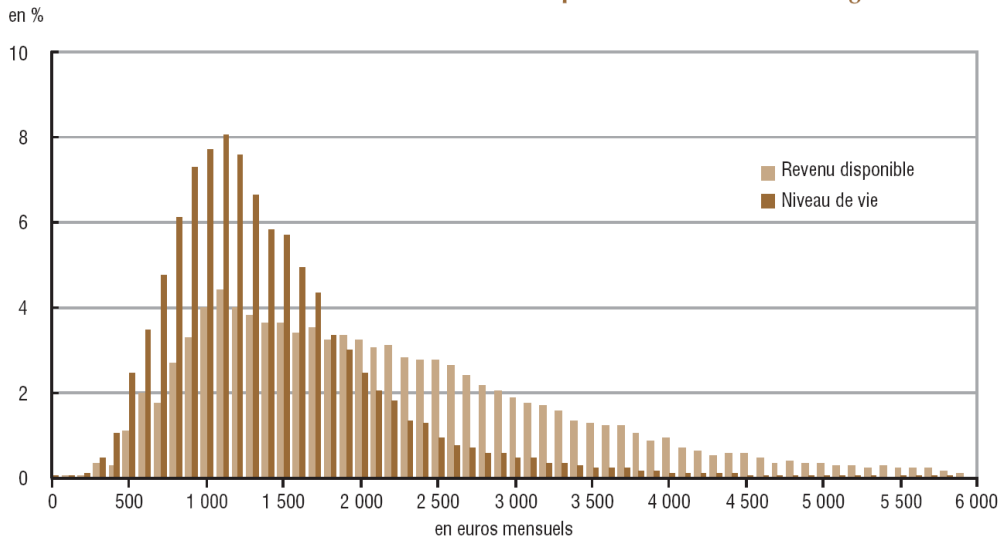
Pour mettre en évidence les caractéristiques de dispersion de la série statistique, on détermine la médiane, les quartiles, les déciles, certains centiles etc ...

Pour cela, on établit le tableau des fréquences cumulées croissantes ou des effectifs cumulés croissants (obtenu directement à l'aide de logiciels lorsqu'il y a de nombreuses données).

Puis on construit un graphique pour faciliter l'interprétation

graphique n°1

2. Distributions des niveaux de vie et du revenu disponible mensuel des ménages en 2004



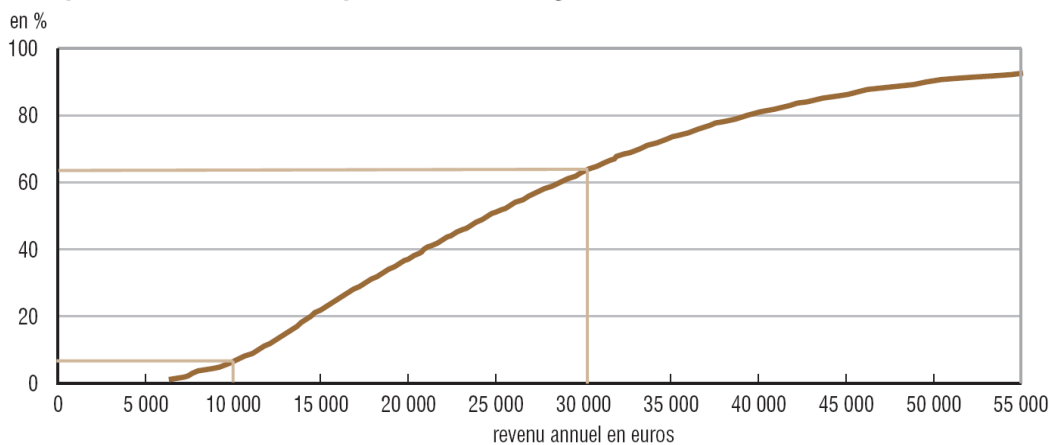
Note : au-delà de 6000 euros, les distributions ne sont pas représentées.

Lecture : le pas de l'histogramme est de 100 euros ; la hauteur de la barre de coordonnées n en abscisse est donc égale à la proportion de revenus ou niveaux de vie compris entre n et $n+100$ euros : ainsi 5,7 % des individus ont un niveau de vie mensuel compris entre 1 500 euros et 1 600 euros et 3,6 % des ménages ont un revenu disponible compris entre 1 500 euros et 1 600 euros.

Source : Insee-DGI, enquête Revenus fiscaux 2004.

graphique n°2

1. Répartition des revenus disponibles des ménages en 2004



Champ : ménages dont le revenu déclaré au fisc est positif ou nul et dont la personne de référence n'est pas étudiante.

Lecture : 57 % des ménages ont en 2004 un revenu disponible annuel compris entre 10 000 et 30 000 euros.

Source : enquête Revenus fiscaux 2004, Insee-DGI.

Déterminer graphiquement la médiane de la série puis interpréter en terme d'aire sur le graphique n°1

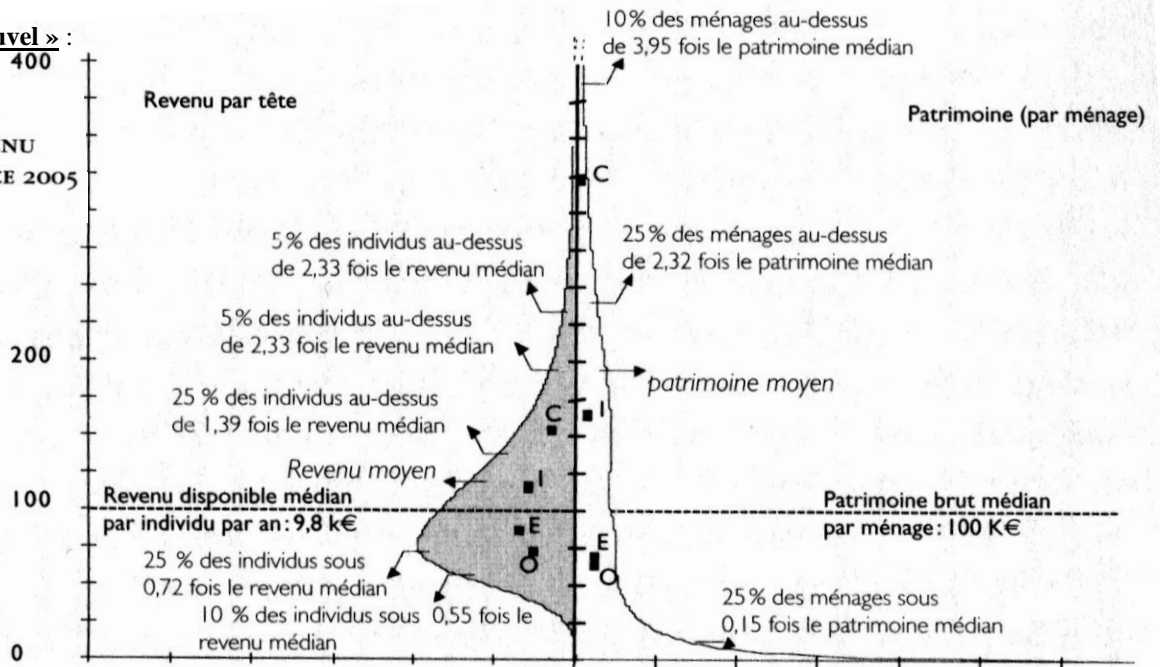
Déterminer la proportion des ménages ayant un revenu disponible compris entre 15 000 € et 35 000 €, interpréter en terme d'aire sur le graphique n°1.

II- Différents indicateurs utilisés en SES:

Pour observer la dispersion de cette série statistique et pouvoir comparer avec d'autres (soit avec des séries de revenus disponibles d'autres pays ou avec des séries d'autres dates ou de comparer deux types de grandeurs à la même date et dans le même pays), on peut utiliser plusieurs indicateurs : $Q3-Q1$, $D9/D1$, construction des courbes de Lorenz associées et le coefficient de Gini (voir TD n°4 sur la construction de telles courbes), mais aussi l'analyse de « strobiloïdes » construits avec des séries médianisées (c'est-à-dire avec des valeurs proportionnelles en divisant les valeurs de la série par la médiane).

Le strobiloïde dit de « Chauvel » :

**Graphique 1
LA RÉPARTITION DU REVENU
ET DU PATRIMOINE EN FRANCE 2005
(strobiloïdes)**



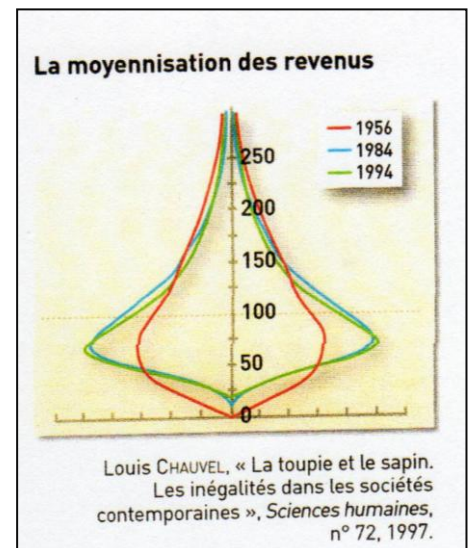
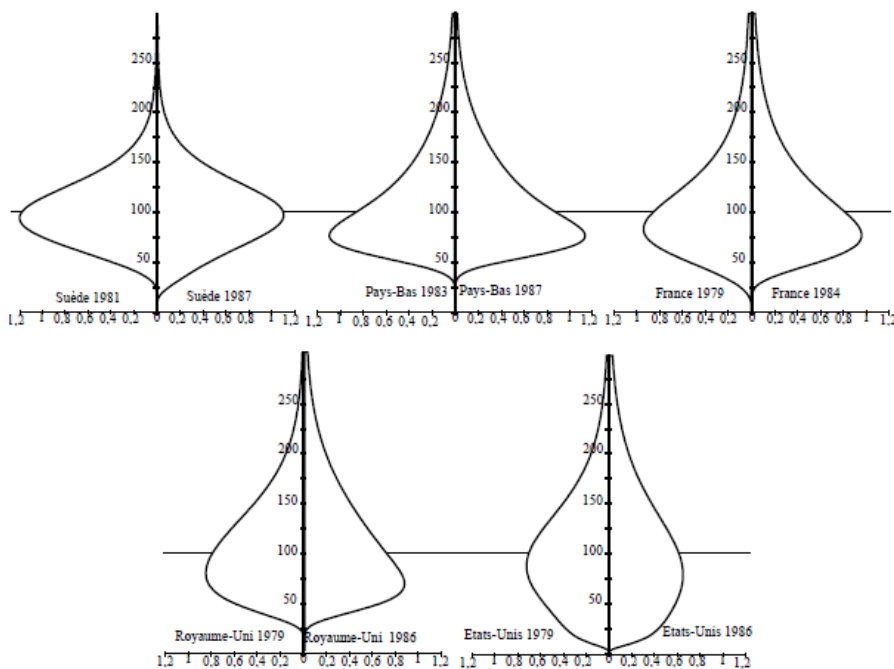
Note : le strobiloïde représente la forme de la pyramide sociale correspondant à la répartition du revenu ou du patrimoine. La courbe (zones grisées) est d'autant plus large qu'une part importante de la population est située précisément à ce niveau. Si 100 représente la médiane des revenus, le large renflement au centre du strobiloïde permet de révéler une forte classe moyenne située à égale distance des extrêmes. Du côté des patrimoines, au contraire, il n'existe pas de classe moyenne dans la mesure où la population est largement étirée entre l'extrême opulence et l'extrême dénuement en termes d'accumulation. Les revenus sont exprimés en francs par tête dans le ménage : l'individu médian se trouve dans un ménage où le revenu annuel est de 9,8 K€ par individu. Le patrimoine est exprimé par ménage. Pour le revenu, C, I, E et O représentent la médiane du revenu respectivement des cadres, des professions intermédiaires, des employés et des ouvriers. Pour le patrimoine, ce sont les moyennes, les médianes n'ayant pas été calculées dans les publications de l'INSEE.

Pour autant, entre la moitié et le double du revenu médian, on compte 80 % de la population. Entre les quatre cinquièmes du revenu médian (médiane -20 %) et les cinq quarts (médiane +25 %), 34 %. Alors que les titulaires des plus bas revenus se raréfient rapidement à mesure que l'on descend en deçà de 75 % du revenu médian, tout comme les plus riches lorsque l'on s'élève dans l'échelle des revenus, nous avons à ce niveau intermédiaire de la répartition du revenu une certaine accumulation de la population : la « classe de revenu médian ».

Comparaison avec ceux d'autres pays et évolution en France

FONDATION NATIONALE DES SCIENCES POLITIQUES

OBSERVATOIRE FRANCAIS DES CONJONCTURES ECONOMIQUES



Louis CHAUVEL, « La toupie et le sapin. Les inégalités dans les sociétés contemporaines », *Sciences humaines*, n° 72, 1997.

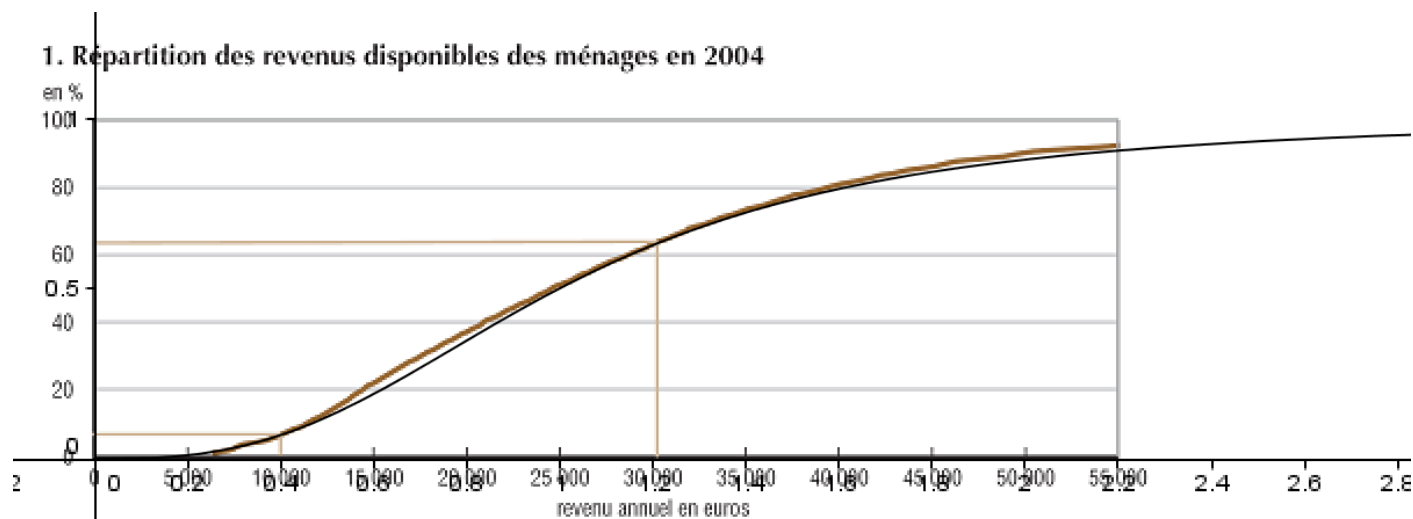
III- Modélisation mathématique du strobiloïde des revenus

L'allure des séries des revenus a été l'objet d'étude de mathématiciens afin d'en dégager des indicateurs. Par exemple, Pareto au XIXème mais il n'avait que des données tronquées (difficulté d'obtenir des données), puis Champernown etc Ces études débouchèrent sur une étude mathématique générale des séries statistiques.

1 – **Une courbe de Champernown du type** : $F(R) = 1 - \frac{1}{1+R^a}$

Prenons par exemple une courbe de Champernown du type $F(R) = 1 - \frac{1}{1+R^a}$ pour modéliser le graphique n°2, c'est-à-dire celui représentant la répartition des **revenus disponibles selon les fréquences cumulées**, avec R le revenu médianisé, c'est-à-dire que la valeur de la médiane est ramenée à 1 (on divise toutes les données par la valeur de la médiane).

A l'aide du logiciel geogebra on observe qu'une valeur de a possible est environ : 2,86



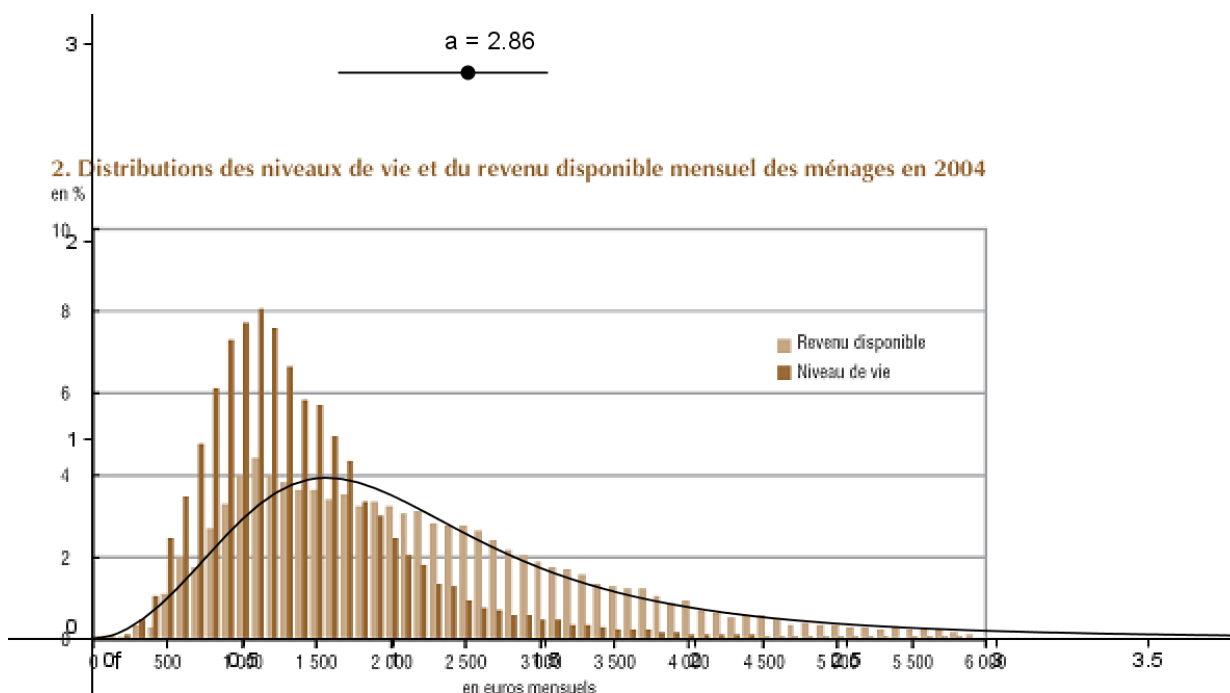
Cette courbe vérifie $F(0)=0$ et $\lim_{R \rightarrow +\infty} F(R) = 1$, c'est-à-dire que plus R est grand, plus F(R) se rapproche de 1, c'est-à-dire de 100% des effectifs !

Cette courbe représente la **fonction de répartition des revenus**, elle approche celle de la répartition des fréquences cumulées des revenus.

2- Lien avec la dérivée :

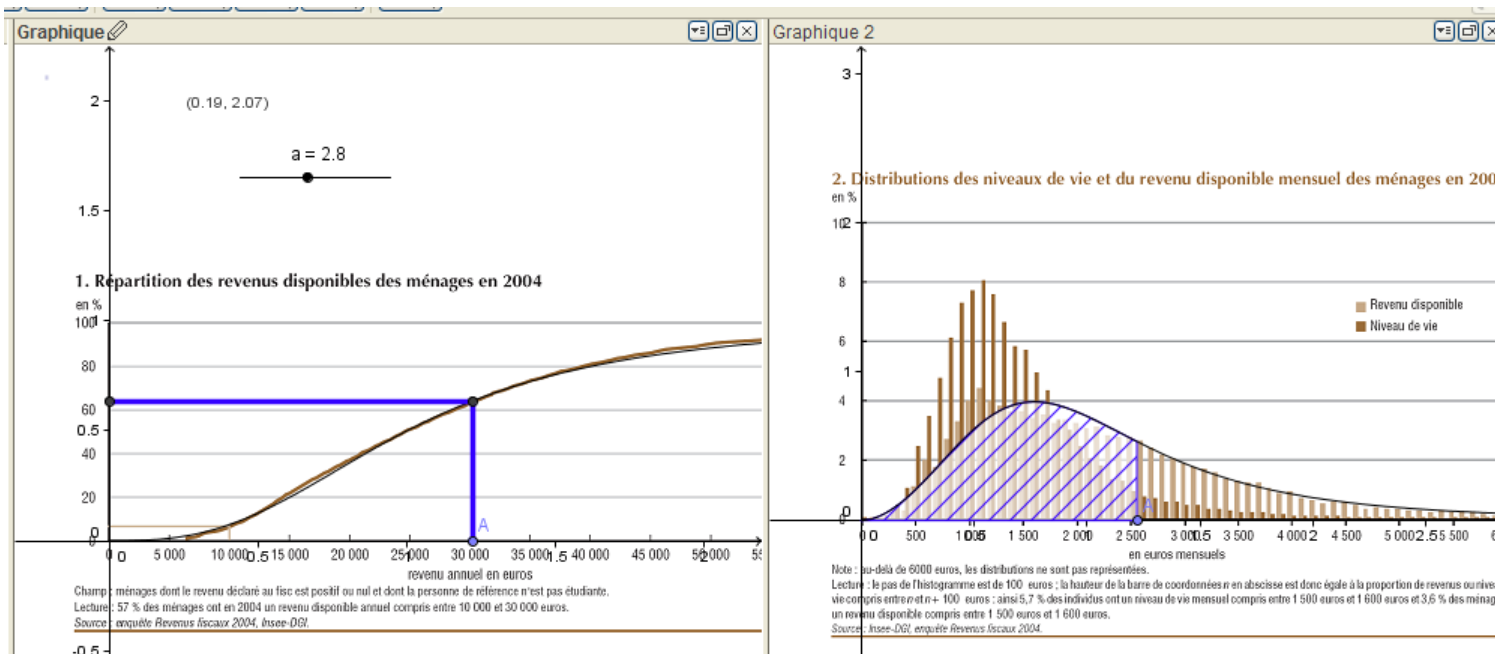
La courbe du graphique n°1 correspond à la **fonction densité des revenus** et peut être modélisée par la dérivée de la fonction F :

$$F'(R) =$$



L'aire située sous la courbe est bien égale à 1 c'est-à-dire à 100% des revenus (voir interprétations graphiques précédentes où les fréquences (ou les effectifs) sont interprétées en termes d'aire) et c'est pour cette raison que les unités ne sont pas respectées sur l'axe des ordonnées (aire de tranches de revenus de largeur infiniment petite).

On retrouve ce résultat par le « calcul intégral » en interprétant le lien entre les deux courbes avec la notion d'aire :



$$F(r) = \int_0^r F'(x)dx = p(x < r) ,$$

F(r) correspond à la proportion des revenus disponibles inférieurs à r, c'est-à-dire à l'aire située entre la courbe représentant F', l'axe des abscisses et les deux droites verticales d'équations x = 0 et x = r.

Cette notation avec le symbole \int ressemble au S de la somme des rectangles de l'histogramme avec une largeur considérée comme infiniment petite.

On dit : « intégrale de la fonction F' de 0 à r »

Remarque : $\lim_{r \rightarrow +\infty} \int_0^r F'(x)dx = \lim_{r \rightarrow +\infty} F(r) - F(0) = 1 - 0 = 1$ noté $\int_0^{+\infty} F'(x)dx = 1$

Interprétation de ces symboles : l'aire située sous la courbe représentant F' est égale à 1 pour r suffisamment grand.

Attention : la notion de « limite » n'est plus au programme de terminale ES, mais elle sera étudiée dans vos études dans le supérieur (c'est une notion mathématique assez intuitive !)

3- Modélisation par des courbes plus compliquées :

La modélisation des courbes de revenus a intéressé le sociologue Louis Chauvel pour en dégager des paramètres qui pourraient permettre des comparaisons fines et exploitables en termes d'aide à la décision sur les politiques de réduction des inégalités. Il utilise des fonctions plus compliquées que celles de Champernown avec trois paramètres permettant de prendre en compte les bas de la courbe du strobiloïde (les plus bas revenus) et les hauts (les hauts revenus) .

Voir animation avec geogebra.

6 La réduction des inégalités de revenus

orsque l'on aborde la dynamique historique, trois résultats sont importants. D'abord, les Trente Glorieuses (1945-1975) apparaissent comme une période extraordinaire d'enrichissement du salariat : alors qu'en France, le pouvoir d'achat du salaire réel a augmenté d'environ 0,5 % par an dans les années 1950, une année moyenne de la période antérieure à 1975 s'accompagnait d'une hausse de 3,5 %. Ensuite, l'écart entre cadres et ouvriers nous prenions les employés, le rapport serait presque identique) a augmenté après 1968 (le cadre gagnait en moyenne 4 fois plus que l'ouvrier en 1968, et 2,7 en 1984) ; un mouvement considérable de rapprochement

des salaires entre qualifiés et routiniers a donc eu lieu après 1968 ; ce mouvement résulte notamment d'un renouvellement fort de la population des cadres dans les années 1970 avec l'arrivée précoce et massive des jeunes générations diplômées du baby-boom. Enfin, depuis cette date, l'écart est à peu près stable (2,5 en 1998). La fin des Trente Glorieuses a donc été marquée par une dynamique doublement favorable à la classe ouvrière (et aux employés) : les écarts se réduisent dans un contexte d'enrichissement rapide et partagé.

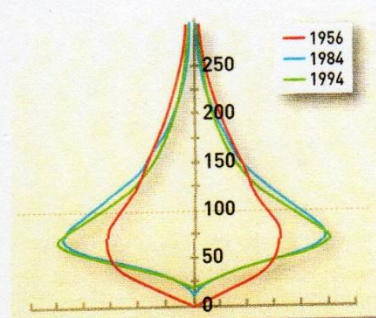
Louis CHAUVEL,
« Le retour des classes sociales ? »
Revue de l'OFCE, n° 79, octobre 2001.

Strobiloïde

est un graphique de répartition en forme de toupie (en grec, *strobilos* signifie toupie). Il se construit en mettant en ordonnées le revenu exprimé sous forme d'indice, et en abscisses la population concernée par ce revenu. La partie gauche et la partie droite du graphique sont symétriques de part et d'autre du revenu médian.

Le point 0 et 50 % du revenu médian se situe le groupe des plus pauvres. Le point 50 et 200 % du revenu médian se situent les catégories moyennes. Au-delà de 200 % se trouve le groupe des plus riches.

La moyennisation des revenus



Louis CHAUVEL, « La toupie et le sapin. Les inégalités dans les sociétés contemporaines », *Sciences humaines*, n° 72, 1997.

QUESTIONS

- 1 Recenser** : quels sont les éléments qui attestent d'une réduction des inégalités ?
- 2 Déduire** : comme le pouvoir d'achat a crû de 3,5 % par an en moyenne entre 1945 et 1975, par combien a-t-il été multiplié sur la période ?
- 3 Décrire** : d'après le strobiloïde, quels sont les signes d'une moyennisation de la société et d'une réduction des inégalités entre 1956 et 1984 ?

Pour la partie modélisation

- 1) **strobiloiseetrepartition1.jpg.ggb** : faire bouger le curseur « a » et voir l'évolution des deux courbes sur les deux fenêtres (F et sa dérivée).
- 2) **strobiloetvertical1.ggb** : pour redresser le strobilo verticalement, puis faire bouger « a »
- 3) **strobiloetverticalcomparaison1.ggb** : le même mais comparaison dans une fenêtre avec les strobilo d'autres pays : bouger « a », d'où l'importance d'améliorer l'ajustement des courbes.
- 4) **strobiloiseetrepartitionhachure1.jpg.ggb** : met en évidence la notion de fonction de répartition et fonction densité : Faire bouger le point A bleu pour mettre en correspondance l'aire sous le strobilo et la fonction de répartition (on observe l'égalité de l'aire et de la valeur de F dans la fenêtre algèbre) ... avec ou sans images de fond des graphique ?
- 5) **strobiloetverticalcomparaisonchauvel1.ggb** : amélioration avec les fonctions de modélisation de Chauvel, faire bouger les a , b et c . Les valeurs remarquables des différents pays sont données dans le tableau qui doit apparaître quand on demande de faire apparaître l'image 2.